

Soft Covering with High Probability

Paul Cuff
Princeton University

Abstract—Wyner’s soft-covering lemma is the central analysis step for achievability proofs of information theoretic security, resolvability, and channel synthesis. It can also be used for simple achievability proofs in lossy source coding. This work sharpens the claim of soft-covering by moving away from an expected value analysis. Instead, a random codebook is shown to achieve the soft-covering phenomenon with high probability. The probability of failure is super-exponentially small in the block-length, enabling many applications through the union bound. This work gives bounds for both the exponential decay rate of total variation and the second-order codebook rate for soft covering.

I. SOFT COVERING

Soft covering of a distribution by a codebook is a concept that was introduced by Wyner [1, Theorem 6.3]. He developed this tool for the purpose of proving achievability in his work on the common information of two random variables. Coincidentally, the most prevalent current application of soft covering is for security proofs in wiretap channels (e.g. [2]), which he also introduced that same year in [3] but apparently did not see how soft covering applied.

We will focus exclusively on the memoryless case, as did Wyner. Given a channel $Q_{Y|X}$ and an input distribution Q_X , let the output distribution be Q_Y . Also, let the n -fold memoryless extensions of these be denoted $Q_{Y^n|X^n}$, Q_{X^n} , and Q_{Y^n} .

Wyner’s soft-covering lemma says that the distribution induced by selecting a X^n sequence at random from a codebook of sequences and passing it through the memoryless channel $Q_{Y^n|X^n}$ will be a good approximation of Q_{Y^n} in the limit of large n as long as the codebook is of size greater than 2^{nR} where $R > I(X; Y)$. In fact, the codebook can be chosen quite carelessly—by random codebook construction, drawing each sequence independently from the distribution Q_{X^n} .

Some illustrations of this phenomenon can be found in Figures 1-6. Here we demonstrate the synthesis of a Gaussian distribution by applying additive Gaussian noise to a codebook. The solid curve in Figures 1, 2, 4, and 5 is the desired output distribution, while the dashed curves are the approximations induced by the codebooks. Along the bottom of each graphic the codewords themselves are illustrated with an ‘x.’ In Figures 3 and 6 for the 2-dimensional case, only the distributions induced by the codewords are shown in gray-scale.

The signal-to-noise-ratio is 15 for these examples, meaning that the variance of the desired output distribution is 16 times that of the additive noise. This gives a mutual information of 2 bits. Accordingly, if the codebook size is b^n , where $b > 4$ and n is the dimension, then the output distribution should become a very close match as the dimension increases. We show two cases: $b = 5$ (i.e. rate of $\log 5$ per channel use) and $b = 32$ (i.e. rate of 5 bits per channel use). Both rates are sufficient asymptotically for soft covering.

In Figure 1 we see that five randomly chosen codewords does a poor job of approximating the desired output distribution. Figure 3 shows the 2-dimensional version of the $b = 5$ example. It’s still a poor approximation. On the other hand, the $b = 32$ example shown in Figure 4 and Figure 6 begins to look quite good already in two dimensions. The benefit of the increased dimension seems apparent, as the distribution is able to have a smoother appearance. This same benefit will ultimately occur in the $b = 5$ case as well, but it will require a higher dimension to manifest itself.

One might also like to consider how good the approximation can be if the codebook is chosen carefully rather than at random. Figure 2 and Figure 5 show this for the two cases in one dimension. We see that while five points is not enough for a great approximation, it can do about as well as 32 randomly chosen codewords. Also, the 32 carefully placed codewords induce an excellent approximation. The study of the best codebooks was performed under the name “resolvability” in [4]. Even though careful placement obviously helps in the cases illustrated, the effect wears off in high dimensions, and you still cannot use a codebook rate $R < I(X; Y)$. Soft covering theorems, as described herein, exclusively consider random codebooks. The intended applications are coding theorems in communication settings where often random codebook generation is convenient.

II. LITERATURE

The soft-covering lemmas in the literature use a distance metric on distributions (commonly total variation or relative entropy) and claim that the distance between the induced distribution P_{Y^n} and the desired distribution Q_{Y^n} vanishes in expectation over the random selection of the set.¹ In the literature, [4] studies the fundamental limits of soft-covering as “resolvability,” [5] provides rates of exponential convergence, [6] improves the exponents and extends the framework, [7] and [8, Chapter 16] refer to soft-covering simply as “covering” in the quantum context, [9] refers to it as a “sampling lemma” and points out that it holds for the stronger metric of relative entropy, and [10] gives a recent direct proof of the relative entropy result. A covering lemma found in [11] makes a high probability claim similar to this work; however, it is different in that it only applies to what is referred therein as “unweighted” channels. They then use this to make very strong claims about channel synthesis for general channels, even for worst-case channel inputs. That covering lemma is closely related to this at a high level but technically quite different.

Here we give two incarnations of a stronger claim. First, with high probability with respect to the codebook distribution, for any fixed rate $R > I(X; Y)$, the total variation distance

¹Many of the theorems only claim existence of a good codebook, but all of the proofs use expected value to establish existence.

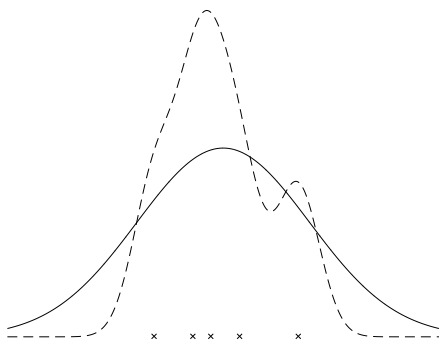


Fig. 1. 5 randomly selected codewords

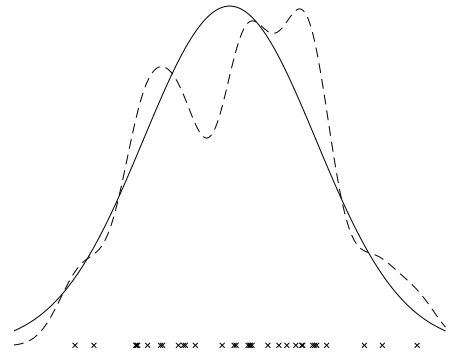


Fig. 4. 32 randomly selected codewords

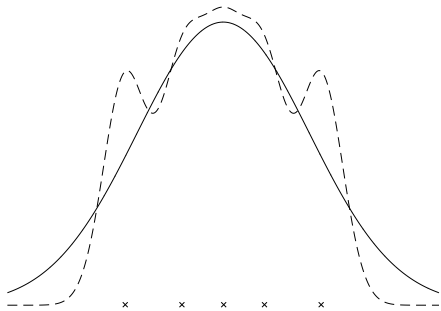


Fig. 2. 5 carefully selected codewords

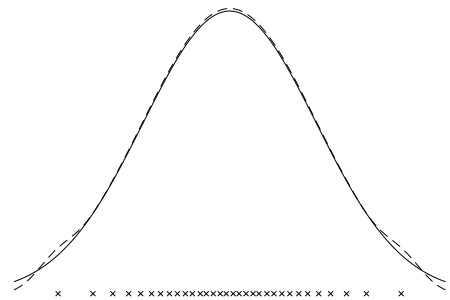


Fig. 5. 32 carefully selected codewords



Fig. 3. $25 = 5^2$ randomly selected codewords in 2 dimensions

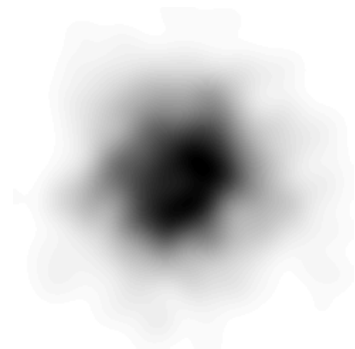


Fig. 6. $1024 = 32^2$ randomly selected codewords in 2 dimensions

will vanish exponentially quickly with the block-length n . Second, if the codebook rate exceeds $I(X; Y)$ by a vanishing amount of order $\frac{1}{\sqrt{n}}$ (referred to as the second-order rate), then the total variation is bounded by a constant with high probability. In both cases, the negligible probability of the random codebook not producing these desired results is super-exponentially small.

Both of the results provided in this work have matching results in the literature, but in a weaker form. For example, the same second-order rate provided in this work was shown in [13] using a bound found in [5]. The difference is that the present work shows that a random codebook will achieve this same phenomenon at the same efficient rates with extremely high probability. Previous results only made claims about the best codebook via expected value arguments. As we will claim in Section IV, the high probability aspect of these results is crucial for solving certain problems that were previously challenging.

The results presented in this paper are highly related to results we presented in [14] and [15]. In this paper we prove the high probability results directly for the total variation metric (instead of relative entropy), which yields tighter results by a factor of two in the exponent than going through Pinsker's inequality. Furthermore, we provide second-order rate results not present in any previous publications.

III. MAIN RESULTS

Let us define precisely the induced distribution. Let $C = \{x^n(m)\}_{m=1}^M$ be the codebook. Then the induced distribution of Y^n , which is a function of the codebook C , is the conditional distribution

$$P_{Y^n|C} = 2^{-nR} \sum_{x^n(m) \in C} Q_{Y^n|X^n=x^n(m)}. \quad (1)$$

The codebook itself is randomly generated, with the code-words mutually independent and distributed according to

$$x^n(m) \sim Q_{X^n} \quad \forall m. \quad (2)$$

We denote the random codebook with calligraphic text, as \mathcal{C} . Thus, $P_{Y^n|C=\mathcal{C}}$ is a random distribution because of the random codebook selection. For notational brevity, we will refer to this simply as $P_{Y^n|\mathcal{C}}$.

The size of the codebook is $M = 2^{nR}$. In the case of Theorem 2, we let the rate vary with n so that it converges down to the asymptotic limit of $I(X; Y)$.

Theorem 1 (Exponential convergence). *For any $Q_X, Q_{Y|X}$, and $R > I(X; Y)$, where X and Y have finite supports \mathcal{X} and \mathcal{Y} , there exists a $\gamma_1 > 0$ and a $\gamma_2 > 0$ such that for n large enough*

$$\mathbb{P}(\|P_{Y^n|\mathcal{C}} - Q_{Y^n}\|_{TV} > e^{-\gamma_1 n}) \leq e^{-e^{\gamma_2 n}}, \quad (3)$$

where $\|\cdot\|_{TV}$ is total variation.

More precisely, for any $n \in \mathbb{N}$ and $\delta \in (0, R - I(X; Y))$,

$$\mathbb{P}(\|P_{Y^n|\mathcal{C}} - Q_{Y^n}\|_{TV} > 3 \cdot 2^{-n\gamma_\delta}) \leq (1 + |\mathcal{Y}|^n) e^{-\frac{1}{3}2^{n\delta}}, \quad (4)$$

where

$$\gamma_\delta = \sup_{\alpha > 1} \frac{\alpha - 1}{2\alpha - 1} (R - \delta - d_\alpha(Q_{X,Y}, Q_X Q_Y)), \quad (5)$$

and $d_\alpha(\cdot, \cdot)$ is the Rényi divergence of order α .

Theorem 2 (Second order rate). *For any $Q_X, Q_{Y|X}$, and $\varepsilon \in (0, 1)$, where X and Y have finite supports \mathcal{X} and \mathcal{Y} , let the rate R vary with n as*

$$R_n = I(X; Y) + \frac{1}{\sqrt{n}} \mathcal{Q}^{-1}(\varepsilon) \sqrt{V} + c \frac{\log n}{n}, \quad (6)$$

where \mathcal{Q} is one minus the standard normal cdf, V is the variance of $\iota_{X;Y}(X; Y)$, and $c > 2$ is arbitrary. Then for any $d < c - 1$ and for n large enough,

$$\mathbb{P}(\|P_{Y^n|\mathcal{C}} - Q_{Y^n}\|_{TV} > \varepsilon) \leq e^{-n^d}. \quad (7)$$

Proof of Theorem 1: We state the proof in terms of arbitrary distributions (not necessarily discrete). When needed, we will specialize to the case that \mathcal{X} and \mathcal{Y} are finite.

Let the Radon-Nikodym derivative between the induced and desired distributions be denoted as

$$D_{\mathcal{C}}(y^n) \triangleq \frac{dP_{Y^n|\mathcal{C}}}{dQ_{Y^n}}(y^n). \quad (8)$$

In the discrete case, this is just a ratio of probability mass functions.

Notice that the total variation of interest, which is a function of the codebook \mathcal{C} , is given by

$$\|P_{Y^n|\mathcal{C}} - Q_{Y^n}\|_{TV} = \frac{1}{2} \int dQ_{Y^n} |D_{\mathcal{C}} - 1| \quad (9)$$

$$= \int dQ_{Y^n} [D_{\mathcal{C}} - 1]_+, \quad (10)$$

where $[z]_+ = \max\{z, 0\}$.

Define the jointly-typical set over x and y sequences by

$$\mathcal{A}_\epsilon \triangleq \left\{ (x^n, y^n) : \frac{1}{n} \log \frac{dQ_{Y^n|X^n=x^n}}{dQ_{Y^n}}(y^n) \leq I(X; Y) + \epsilon \right\}. \quad (11)$$

We split $P_{Y^n|\mathcal{C}}$ into two parts, making use of the indicator function denoted by $\mathbf{1}$. Let $\epsilon > 0$ be arbitrary, to be determined later.

$$P_{\mathcal{C},1} \triangleq 2^{-nR} \sum_{x^n(m) \in \mathcal{C}} Q_{Y^n|X^n=x^n(m)} \mathbf{1}_{(Y^n, x^n(m)) \in \mathcal{A}_\epsilon}, \quad (12)$$

$$P_{\mathcal{C},2} \triangleq 2^{-nR} \sum_{x^n(m) \in \mathcal{C}} Q_{Y^n|X^n=x^n(m)} \mathbf{1}_{(Y^n, x^n(m)) \notin \mathcal{A}_\epsilon}. \quad (13)$$

The measures $P_{\mathcal{C},1}$ and $P_{\mathcal{C},2}$ on the space \mathcal{Y}^n are not probability measures, but $P_{\mathcal{C},1} + P_{\mathcal{C},2} = P_{Y^n|\mathcal{C}}$ for each codebook \mathcal{C} .

Let us also split $D_{\mathcal{C}}$ into two parts:

$$D_{\mathcal{C},1}(y^n) \triangleq \frac{dP_{\mathcal{C},1}}{dQ_{Y^n}}(y^n), \quad (14)$$

$$D_{\mathcal{C},2}(y^n) \triangleq \frac{dP_{\mathcal{C},2}}{dQ_{Y^n}}(y^n). \quad (15)$$

This allows us also to bound the total variation by a sum of two terms:

$$\|P_{Y^n|C} - Q_{Y^n}\|_{TV} \leq \int dQ_{Y^n}[D_{C,1} - 1]_+ + \int dQ_{Y^n} D_{C,2} \quad (16)$$

$$= \int dQ_{Y^n}[D_{C,1} - 1]_+ + \int dP_{C,2}. \quad (17)$$

Notice that $P_{C,1}$ will usually contain almost all of the probability. That is, denoting the complement of \mathcal{A}_ϵ as $\overline{\mathcal{A}_\epsilon}$,

$$\int dP_{C,2} = 1 - \int dP_{C,1} \quad (18)$$

$$= 2^{-nR} \sum_{x^n(m) \in C} \mathbb{P}_Q(\overline{\mathcal{A}_\epsilon} \mid X^n = x^n(m, C)). \quad (19)$$

This is an average of exponentially many i.i.d. random variables bounded between 0 and 1. Furthermore, the expected value of each one is the exponentially small probability of correlated sequences being atypical:

$$\mathbb{E} \mathbb{P}_Q(\overline{\mathcal{A}_\epsilon} \mid X^n = x^n(m, C)) = \mathbb{P}_Q(\overline{\mathcal{A}_\epsilon}) \quad (20)$$

$$\leq 2^{-\beta n}, \quad (21)$$

where

$$\beta = (\alpha - 1)(I(X; Y) + \epsilon - d_\alpha(Q_{X,Y}, Q_X Q_Y)) \quad (22)$$

for any $\alpha > 1$, where $d_\alpha(\cdot, \cdot)$ is the Rényi divergence of order α . Here we use the finiteness of \mathcal{X} and \mathcal{Y} to assure that the Rényi divergence is finite and continuous for all α , which likewise assures a choice of α can be found to give a positive value of β if ϵ is small enough. We use units of bits for mutual information and Rényi divergence to coincide with the base two expression of rate.

Therefore, the Chernoff bound assures that $\int dP_{C,2}$ is exponentially small. That is,

$$\mathbb{P}\left(\int dP_{C,2} \geq 2 \cdot 2^{-\beta n}\right) \leq e^{-\frac{1}{3}2^{n(R-\beta)}}. \quad (23)$$

Similarly, $D_{C,1}$ is an average of exponentially many i.i.d. and uniformly bounded functions, each one determined by one sequence in the codebook:

$$D_{C,1}(y^n) = 2^{-nR} \sum_{x^n(m) \in C} \frac{dQ_{Y^n|X^n=x^n(m)}(y^n)}{dQ_{Y^n}} \mathbf{1}_{(y^n, x^n(m)) \in \mathcal{A}_\epsilon} \quad (24)$$

For every term in the average, the indicator function bounds the value to be between 0 and $2^{nI(X;Y)+n\epsilon}$. The expected value of each term with respect to the codebook is bounded above by one, which is observed by removing the indicator function. Therefore, the Chernoff bound assures that $D_{C,1}$ is exponentially close to one for every y^n . For any $\beta_2 > 0$:

$$\mathbb{P}(D_{C,1}(y^n) \geq 1 + 2^{-\beta_2 n}) \leq e^{-\frac{1}{3}2^{n(R-I(X;Y)-\epsilon-2\beta_2)}} \quad \forall y^n. \quad (25)$$

This use of the Chernoff bound has been used before for a soft-covering lemma in the proof of Lemma 9 of [7].

At this point we will use the fact that \mathcal{Y} is a finite set. We use the union bound applied to (21) and (25), taking advantage of the fact that the space \mathcal{Y}^n is only exponentially large. Let \mathcal{S} be the set of codebooks such that the following are true:

$$\int dP_{C,2} < 2 \cdot 2^{-\beta n}, \quad (26)$$

$$D_{C,1}(y^n) < 1 + 2^{-\beta_2 n} \quad \forall y^n \in \mathcal{Y}^n. \quad (27)$$

We see that the probability of not being in \mathcal{S} is doubly exponentially small:

$$\mathbb{P}(C \notin \mathcal{S}) \leq e^{-\frac{1}{3}2^{n(R-\beta)}} + |\mathcal{Y}|^n e^{-\frac{1}{3}2^{n(R-I(X;Y)-\epsilon-2\beta_2)}}. \quad (28)$$

What remains is to show that for every codebook in \mathcal{S} , the total variation is exponentially small. From (17) it follows that

$$\|P_{Y^n|C} - Q_{Y^n}\|_{TV} \leq 2 \cdot 2^{-\beta n} + 2^{-\beta_2 n}. \quad (29)$$

Finally, we carefully select ϵ , β_1 and β_2 to give the tightest exponential bound, in terms of δ from the theorem statement:

$$\epsilon_{\alpha,\delta} = \frac{\frac{1}{2}(R-\delta) + (\alpha-1)d_\alpha(Q_{X,Y}, Q_X Q_Y)}{\frac{1}{2} + (\alpha-1)} - I(X; Y), \quad (30)$$

$$\beta_2 = \beta. \quad (31)$$

This gives the desired result. \blacksquare

Proof of Theorem 2: For the analysis of the second-order rate, we use many of the same steps as the proof of Theorem 1. Assume all of the same definitions.

The key difference is the bound on $\mathbb{P}_Q(\overline{\mathcal{A}_\epsilon})$ found in (21). Instead of using the Chernoff bound we will use the Berry-Esseen theorem.

$$\mathbb{P}_Q(\overline{\mathcal{A}_\epsilon}) \leq \mathcal{Q}\left(\frac{\epsilon\sqrt{n}}{\sqrt{V}}\right) + \frac{\rho}{V^{3/2}\sqrt{n}}, \quad (32)$$

where $\rho = \mathbb{E}|\imath_{X;Y}(X; Y) - I(X; Y)|^3 \leq \infty$ because \mathcal{X} and \mathcal{Y} are finite.

Now choose $r \in (0, c-d-1)$, where c and d are from the theorem statement, and let

$$\epsilon = \frac{1}{\sqrt{n}} \mathcal{Q}^{-1}(\varepsilon) \sqrt{V} + r \frac{\log n}{n}. \quad (33)$$

The bound on $\mathbb{P}_Q(\overline{\mathcal{A}_\epsilon})$ becomes

$$\mu_n \triangleq \mathcal{Q}\left(\mathcal{Q}(\varepsilon) + \frac{r}{\sqrt{V}} \frac{\log n}{\sqrt{n}}\right) + \frac{\rho}{V^{3/2}\sqrt{n}}. \quad (34)$$

Now, in a step analogous to (23), again using the Chernoff bound,

$$\mathbb{P}\left(\int dP_{C,2} \geq \mu_n \left(1 + \frac{1}{\sqrt{n}}\right)\right) \leq e^{-\frac{\mu_n}{3n} 2^{nR}}. \quad (35)$$

Also, in a step analogous to (25), we use the Chernoff bound to obtain the following:

$$\mathbb{P}\left(D_{C,1}(y^n) \geq 1 + \frac{1}{\sqrt{n}}\right) \leq e^{-\frac{1}{3n} 2^{n(R-I(X;Y)-\epsilon)}} \quad \forall y^n \quad (36)$$

$$= e^{-\frac{1}{3n} 2^{n((c-r)\frac{\log n}{n})}} \quad (37)$$

$$= e^{-\frac{1}{3} n^{c-r-1}}. \quad (38)$$

At this point we will use the fact that \mathcal{Y} is a finite set to apply the union bound. Let $\bar{\mathcal{S}}$ be the set of codebooks such that the following are true:

$$\int dP_{\mathcal{C},2} < \mu_n \left(1 + \frac{1}{\sqrt{n}}\right), \quad (39)$$

$$D_{\mathcal{C},1}(y^n) < 1 + \frac{1}{\sqrt{n}} \quad \forall y^n \in \mathcal{Y}^n. \quad (40)$$

The probability of not being in $\bar{\mathcal{S}}$ is super-exponentially small:

$$\mathbb{P}(\mathcal{C} \notin \bar{\mathcal{S}}) \leq e^{-\frac{\mu_n}{3n} 2^{nR}} + |\mathcal{Y}|^n e^{-\frac{1}{3} n^{c-r-1}}. \quad (41)$$

Notice that μ_n converges to ε , so the above probability is dominated by the second term. Since $d < c - r - 1$, this establishes the probability statement of the theorem.

What remains is to show that for every codebook in $\bar{\mathcal{S}}$, the total variation is eventually less than ε . From (17) it follows that

$$\|P_{Y^n|\mathcal{C}} - Q_{Y^n}\|_{TV} \leq \mu_n \left(1 + \frac{1}{\sqrt{n}}\right) + \frac{1}{\sqrt{n}}. \quad (42)$$

As observed previously, μ_n converges to ε . Furthermore, it converges from below and the deviation is of order $\frac{\log n}{\sqrt{n}}$, which dominates the other terms in the total variation bound. ■

IV. APPLICATIONS

As stated in [14], these stronger versions of Wyner's soft-covering lemma have important applications, particularly to information theoretic security. The main advantage of these theorems come from the union bound.

The usual random coding argument for information theory uses a randomly generated codebook until the final steps of the achievability proof. In these final steps, it is claimed that there exists a good codebook based on the analysis. This can be done by analyzing the expected value of the performance for the random ensemble and claiming that at least one codebook is as good as the expected value. Alternatively, one can make the argument based on the probability that the randomly generated codebook has a good performance. If that probability is greater than zero, then there is at least one good codebook. The second approach can be advantageous when performance is not captured by one scalar value that is easily analyzed—for example, if “good” performance involves a collection of constraints.

These stronger soft-covering theorems give a very strong assurance that soft-covering will hold. Even if the codebook needs to satisfy exponentially many constraints related to soft-covering, the union bound will yield the claim that a codebook exists which satisfies them all simultaneously. Indeed, if you ran the soft-covering experiment exponentially many times, regardless of how the codebooks are correlated from one experiment to the next, the probability of seeing even one fail is still super-exponentially small.

We have shown in [15] that high probability theorems for soft covering can be used to tackle previously challenging problems in secure communication where precisely this need arises. That work demonstrates that in the wiretap channel of type II [12], where an adversary can influence the channel with an exponential number of possible actions, a random

codebook will achieve secrecy for all of them simultaneously. Furthermore, “semantic security,” which is a very practical notion of secrecy but is stronger than the secrecy typically guaranteed in information theory, requires security to hold even for the most distinguishable pair of messages. This level of secrecy is shown to be achieved, again using the union bound and the super-exponential assurance of soft covering.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation (grant CCF-1350595) and the Air Force Office of Scientific Research (grant FA9550-15-1-0180).

REFERENCES

- [1] A. Wyner, “The common information of two dependent random variables,” *IEEE Trans. Inf. Theory*, 21(2): 163-79, March 1975.
- [2] M. Bloch and N. Laneman, “Strong secrecy from channel resolvability,” *IEEE Trans. Inf. Theory*, 59(12): 8077-8098, Dec. 2013.
- [3] A. Wyner, “The wire-tap channel,” *Bell Systems Technical Journal*, 54(8): 1334-87, Oct. 1975.
- [4] T. Han and S. Verdú, “Approximation theory of output statistics,” *IEEE Trans. Inf. Theory*, 39(3): 752-72, May 1993.
- [5] M. Hayashi, “General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel,” *IEEE Trans. Inf. Theory*, 52(4): 1562-75, April 2006.
- [6] P. Cuff, “Distributed channel synthesis,” *IEEE Trans. Inf. Theory*, 59(11): 7071-96, Nov. 2013.
- [7] R. Ahlswede and A. Winter, “Strong converse for identification via quantum channels,” *IEEE Trans. Inf. Theory*, 48(3): 569-79, March 2002.
- [8] M. Wilde, “Quantum information theory,” *Cambridge University Press*, 2013.
- [9] A. Winter, “Secret, public and quantum correlation cost of triples of random variables,” *Proc. of IEEE Int'l. Symp. Inf. Theory*, Sept. 2005.
- [10] J. Hou and G. Kramer, “Effective secrecy: Reliability, confusion and stealth,” *Proc. of IEEE Int'l. Symp. Inf. Theory*, July 2014.
- [11] C. Bennett, I. Devetak, A. Harrow, P. Shor, A. Winter. “The Quantum Reverse Shannon Theorem and Resource Tradeoffs for Simulating Quantum Channels,” *IEEE Trans. Inf. Theory*, 60(5), 2926-59, May 2014.
- [12] L. Ozarow and A. Wyner, “Wire-tap channel II,” *Bell Systems Technical Journal*, 63(10): 2135-57, Dec. 1984.
- [13] S. Watanabe and M. Hayashi, “Strong converse and second-order asymptotics of channel resolvability,” *Proc. IEEE Int'l. Symp. Inf. Theory*, July, 2014.
- [14] P. Cuff, “A Stronger Soft-Covering Lemma and Applications,” *Proc. CNS Workshop on Physical-layer Methods for Wireless Security*, Sept. 2015.
- [15] Z. Goldfeld, P. Cuff, H. Permuter. “Semantic-Security Capacity for Wiretap Channels of Type II,” *CoRR:abs/1509.03619*, 2015.